(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

B5

(19) **World Intellectual Property Organization**
International Bureau

(43) **International Publication Date**
25 July 2002 (25.07.2002)

PCT

(10) **International Publication Number**
WO 02/057453 A2

(51) **International Patent Classification⁷:** C12N 15/12, C07K 14/485, 16/22, C12Q 1/68, G01N 33/50, A61K 31/7088, 38/18

(21) **International Application Number:** PCT/US01/50331

(22) **International Filing Date:** 19 December 2001 (19.12.2001)

(25) **Filing Language:** English

(26) **Publication Language:** English

(30) **Priority Data:**

60/265,704	19 December 2000 (19.12.2000)	US
60/257,314	20 December 2000 (20.12.2000)	US
60/288,153	2 May 2001 (02.05.2001)	US
60/294,075	29 May 2001 (29.05.2001)	US
60/307,506	24 July 2001 (24.07.2001)	US
60/311,613	10 August 2001 (10.08.2001)	US
60/311,590	10 August 2001 (10.08.2001)	US
60/315,617	29 August 2001 (29.08.2001)	US
60/322,358	14 September 2001 (14.09.2001)	US

(63) **Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:**

US	60/311,590 (CON)
Filed on	10 August 2001 (10.08.2001)
US	60/265,704 (CON)
Filed on	19 December 2000 (19.12.2000)
US	60/257,314 (CON)
Filed on	20 December 2000 (20.12.2000)
US	60/311,613 (CON)
Filed on	10 August 2001 (10.08.2001)
US	60/315,617 (CON)
Filed on	29 August 2001 (29.08.2001)
US	60/307,506 (CON)
Filed on	24 July 2001 (24.07.2001)
US	60/322,358 (CON)
Filed on	14 September 2001 (14.09.2001)
US	60/294,075 (CON)
Filed on	29 May 2001 (29.05.2001)
US	60/288,153 (CON)
Filed on	2 May 2001 (02.05.2001)

(71) **Applicant** *(for all designated States except US)*: **CURAGEN CORPORATION** [US/US]; 555 Long Wharf Drive, 11th Floor, New Haven, CT 06511 (US).

(72) **Inventors; and**
(75) **Inventors/Applicants** *(for US only)*: **GANGOLLI, Esha, A.** [IN/US]; 31 Strawberry Hill Road, Madison, CT 06443 (US). **PATTURAJAN, Meera** [IN/US]; 45 Harrison Avenue, Apartment 1C, Branford, CT 06405 (US). **VERNET, Corine, A., M.** [FR/US]; 1739 Foxon Road, Box L6, North Branford, CT 06471 (US). **MALYANKAR, Uriel, M.** [IN/US]; 229 Branford Road, #330, Branford, CT 06405 (US). **KEKUDA, Ramesh** [IN/US]; 168 Lockwood Avenue, Stamford, CT 06902 (US). **STONE, David, J.** [US/US]; 223 Whitehorn Drive, Guilford, CT 06437 (US). **ANDERSON, David** [US/US]; 555 Long Wharf Drive, 11th Floor, New Haven, CT 06511 (US). **SHIMKETS, Richard, A.** [US/US]; 5 Ludian Meadows Drive, Guilford, CT 06437 (US). **BURGESS, Catherine, E.** [US/US]; 90 Carriage Hill Drive, Wethersfield, CT 06109 (US). **ZERHUSEN, Bryan, D.** [US/US]; 337 Monticello Drive, Branford, CT 06405 (US). **LIU, Xiaohong** [CN/US]; 90 Montoya Circle, Branford, CT 06405 (US). **SPYTEK, Kimberly, A.** [US/US]; 28 Court Street #1, New Haven, CT 06511 (US). **CASMAN, Stacie, J.** [US/US]; 17 Peck Street, North Haven, CT 06473 (US). **BOLDOG, Ferenc, L.** [HU/US]; 1687 Hartford Turnpike, North Haven, CT 06473 (US). **SMITHSON, Glennda** [US/US]; 125 Michael Drive, Guilford, CT 06435 (US). **LI, li** [CN/US]; 56 Jerimoth Drive, Branford, CT 06405 (US). **JI, Weizhen** [CN/US]; 3 Business Park Drive, Room 101, Branford, CT 06405 (US).

(74) **Agent: ELRIFI, Ivor, R.**; Mintz, Levin, Cohn, Ferris,, Glovsky and Popeo, P.G., One Fiancial Center, Boston, MA 02111 (US).

(81) **Designated States** *(national)*: AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DE, DK, DM, DZ, EC, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, OM, PH, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) **Designated States** *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZM, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent

[Continued on next page]

(54) Title: POLYPETIDES AND NUCLEIC ACIDS ENCODING SAME

(57) **Abstract:** Disclosed herein are nucleic acid sequences that encode novel polypeptides. Also disclosed are polypeptides encoded by these nucleic acid sequences, and antibodies, which immunospecifically-bind to the polypeptide, as well as derivatives, variants, mutants, or fragments of the aforementioned polypeptide, polynucleotide, or antibody. The invention further discloses therapeutic, diagnostic and research methods for diagnosis, treatment, and prevention of disorders involving any one of these novel human nucleic acids and proteins.

WO 02/057453 A2

(BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).

Published:

— *without international search report and to be republished upon receipt of that report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

polynucleotides" and the corresponding encoded polypeptides are referred to as "NOVX polypeptides" or "NOVX proteins." Unless indicated otherwise, "NOVX" is meant to refer to any of the novel sequences disclosed herein. Table A provides a summary of the NOVX nucleic acids and their encoded polypeptides.

TABLE A. Sequences and Corresponding SEQ ID Numbers

NOVX ASSIGNMENT	Internal Identification	SEQ ID NO (nucleic acid)	SEQ ID NO (polypeptide)	Homology
1	CG55758-01	1	2	SCUBE1-like
2a	CG55724-01	3	4	Adipocyte Complement Related Protein
2b	CG55724-03	5	6	Cq1 TNF-like
2c	CG55724-04	7	8	Cq1 TNF-like
2d	CG55724-06	9	10	Cq1 TNF-like
3	CG50345-01	11	12	β-Adrenergic Receptor Kinase-like
4	CG50301-01	13	14	TENM4-like
5a	CG55764-01	15	16	Out At First-like
5b	CG55764-02	17	18	Out At First-like
6a	CG55704-01	19	20	EphA6-ehk-like
6b	CG55704-03	21	22	EphA6-ehk-like
7	CG94323538	23	24	Glucose Transporter-like
8	CG95545-01	25	26	Type Ia Membrane Sushi-containing domain
9	CG95545-02	27	28	Type Ia Membrane Sushi-containing domain
10a	CG55746-01	29	30	Butyrophilin-like
10b	CG55746-05	31	32	Butyrophilin Precursor B7-DC
11	CG50329-01	33	34	Butyrophilin-like

NOVX nucleic acids and their encoded polypeptides are useful in a variety of applications and contexts. The various NOVX nucleic acids and polypeptides according to the invention are useful as novel members of the protein families according to the presence of domains and sequence relatedness to previously described proteins. Additionally, NOVX nucleic acids and polypeptides can also be used to identify proteins that are members of the family to which the NOVX polypeptides belong.

NOV1 is homologous to an EGF-Related SCUBE1-like family of proteins. Thus, the NOV1 nucleic acids, polypeptides, antibodies and related compounds according to the

section below. The disclosed NOV3 polypeptide has multiple hydrophilic regions, each of which can be used as an immunogen. In one embodiment, a contemplated NOV3 epitope is from about amino acids 20 to 70. In another embodiment, a contemplated NOV3 epitope is from about amino acids 95 to 115. In other specific embodiments, contemplated NOV3 epitopes are from about amino acids 120 to 190, 280 to 300, 305 to 375, 395 to 420, and 415 to 660.

NOV4

A disclosed NOV4 nucleic acid of 8354 nucleotides is set forth as SEQ ID NO:13 (designated CuraGen Acc. No. CG50301-01) encoding a TEN-M4-like protein is shown in Table 4A. An open reading frame was identified beginning with an ATG initiation codon at nucleotides 35-37 and ending with a TAG codon at nucleotides 8342-8344. Putative untranslated regions are indicated by underline.

Table 4A.
NOV4 Polynucleotide
SEQ ID NO:13

GTTTGTGGATGTGGAGGAGCGCGGGCCGGAGGCCATGGACGTGAAGGAGAGGAAGCCTTA	60
CCGCTCGCTGACCCGGCGCCGCGACGCCGAGCGCCGCTACACCAGCTCGTCCGCGGACAG	120
CGAGGAGGGCAAAGCCCCGCAGAAATCGTACAGCTCCAGCGAGACCCTGAAGGCCTACGA	180
CCAGGACGCCCGCCTAGCCTATGGCAGCCGCGTCAAGGACATTGTGCCGCAGGAGGCCGA	240
GGAATTCTGCCGCACAGGTGCCAACTTCACCCTGCGGGAGCTGGGGCTGGAAGAAGTAAC	300
GCCCCCTCACGGGACCCTGTACCGGACAGACATTGGCCTGCCCCAATGCGGCTACTCCAT	360
GGGGGCTGGCTCTGATGCCGACATGGAGGCTGACACGGTGCTGTCCCCTGAGCACCCCGT	420
GCGTCTGTGGGGCCGGAGCACACGGTCAGGGCGCAGCTCCTGCCTGTCCAGCCGGGCCAA	480
TTCCAATCTCACACTCACCGACACCGAGCATGAAAACACTGAGACTGATCATCCGGGCGG	540
CCTGCAGAACCACGCGCGGCTCCGGACGCCGCCGCCGCCGCTCTCGCACGCCCACACCCC	600
CAACCAGCACCACGCGGCCTCCATTAACTCCCTGAACCGGGGCAACTTCACGCCGAGGAG	660
CAACCCCAGCCCGGCCCCCACGGACCACTCGCTCTCCGGAGAGCCCCCTGCCGGCGGCGC	720
CCAGGAGCCTGCCCACGCCCAGGAGAACTGGCTGCTCAACAGCAACATCCCCCTGGAGAC	780
CAGGAACCTAGGCAAGCAGCCATTCCTAGGGACATTGCAGGACAACCTCATTGAGATGGA	840
CATTCTCGGCGCCTCCCGCCATGATGGGGCTTACAGTGACGGGCACTTCCTCTTCAAGCC	900
TGGAGGCACCTCCCCGCTCTTCTGCACCACATCACCAGGGTACCCACTGACGTCCAGCAC	960
AGTGTACTCTCCTCCGCCCCGACCCCTGCCCCGCAGCACCTTCGCCCGGCCGGCCTTTAA	1020
CCTCAAGAAGCCCTCCAAGTACTGTAACTGGAAGTGCGCAGCCCTGAGCGCCATCGTCAT	1080
CTCAGCCACTCTGGTCATCCTGCTGGCATACTTTGTGGCCATGCACCTGTTTGGCCTAAA	1140
CTGGCACCTGCAGCCGATGGAGGGGCAGATGTATGAGATCACGGAGGACACAGCCAGCAG	1200
TTGGCCTGTGCCAACCGACGTCTCCCTATACCCCTCAGGGGGCACTGGCTTAGAGACCCC	1260
TGACAGGAAAGGCAAAGGAACCACAGAAGGAAAGCCCAGTAGTTTCTTTCCAGAGGACAG	1320
TTTCATAGATTCTGGAGAAATTGATGTGGGAAGGCGAGCCTCCCAGAAGATTCCTCCTGG	1380
CACTTTCTGGAGATCTCAAGTGTTCATAGACCATCCTGTGCATCTGAAATTCAATGTGTC	1440
TCTGGGAAAGGCAGCCCTGGTTGGCATTTATGGCAGAAAAGGCCTCCCTCCTTCACATAC	1500
ACAGTTTGACTTTGTGGAGCTGCTGGATGGCAGGAGGCTCCTAACCCAGGAGGCGCGGAG	1560
CCTAGAGGGGACCCCGCGCCAGTCTCGGGGAACTGTGCCCCCCTCCAGCCATGAGACAGG	1620
CTTCATCCAGTATTTGGATTCAGGAATCTGGCACTTGGCTTTTTACAATGACGGAAAGGA	1680
GTCAGAAGTGGTTTCCTTTCTCACCACTGCCATTGAGTCGGTGGATAACTGCCCCAGCAA	1740
CTGCTATGGCAATGGTGACTGCATCTCTGGGACCTGCCACTGCTTCCTGGGTTTCCTGGG	1800
CCCCGACTGTGGCAGAGCCTCCTGCCCCGTGCTCTGTAGCGGAAATGGCCAATACATGAA	1860
AGGCAGATGCTTGTGCCACAGTGGCTGGAAAGGCGCTGAGTGCGATGTGCCCACCAACCA	1920
GTGTATCGATGTGGCCTGCAGCAACCATGGCACCTGCATCACGGGCACCTGCATCTGCAA	1980
CCCTGGCTACAAGGGCGAGAGCTGTGAGGAAGTGGACTGCATGGACCCCACATGTTCAGG	2040
CCGGGGTGTCTGCGTGAGAGGCGAATGCCATTGCTTTGTGGGATGGGGAGGCACCAACTG	2100
CGAGACCCCCAGGGCCACATGCTTAGACCAGTGTTCAGGCCACGGAACCTTCCTCCCGGA	2160
CACCGGGCTTTGCAGCTGTGACCCAAGCTGGACTGGACACGACTGTTCTATCGAGATCTG	2220

```
TGCTGCCGACTGTGGTGGCCATGGCGTGTGCGTAGGGGGCACCTGCCGCTGCGAGGATGG   2280
CTGGATGGGGGCAGCCTGCGACCAGCGGGCCTGCCACCCGCGCTGTGCCGAGCATGGGAC   2340
CTGCCGCGACGGCAAGTGCGAGTGCAGCCCTGGCTGGAATGGCGAACACTGCACCATCGC   2400
TCACTATCTGGATAGGGTAGTTAAAGAGGGTTGCCCTGGGTTGTGCAATGGCAACGGCAG   2460
ATGTACCTTAGACCTGAATGGTTGGCACTGCGTCTGCCAGCTGGGCTGGAGAGGAGCTGG   2520
CTGTGACACTTCCATGGAGACTGCCTGCGGTGACAGCAAAGACAATGATGGAGATGGCCT   2580
GGTGGACTGCATGGACCCTGACTGCTGCCTCCAGCCCCTGTGCCATATCAACCCGCTGTG   2640
CCTTGGCTCCCCTAACCCTCTGGACATCATCCAGGAGACACAGGTCCCTGTGTCACAGCA   2700
GAACCTACACTCCTTCTATGACCGCATCAAGTTCCTCGTGGGCAGGGACAGCACGCACAT   2760
AATCCCCGGGGAGAACCCCTTTGATGGAGGGCATGCTTGTGTTATTCGTGGCCAAGTGAT   2820
GACATCAGATGGAACCCCCCTGGTTGGTGTGAACATCAGTTTTGTCAATAACCCTCTCTT   2880
TGGATATACAATCAGCAGGCAAGATGGCAGCTTTGACTTGGTGACAAATGGCGGCATCTC   2940
CATCATCCTGCGGTTCGAGCGGGCACCTTTCATCACACAGGAGCACACCCTGTGGCTGCC   3000
ATGGGATCGCTTCTTTGTCATGGAAACCATCATCATGAGACATGAGGAGAATGAGATTCC   3060
CAGCTGTGACCTGAGCAATTTTGCCCGCCCCAACCCAGTCGTCTCTCCATCCCCACTGAC   3120
GTCCTTCGCCAGCTCCTGTGCAGAGAAAGGCCCCATTGTGCCGGAAATTCAGGCTTTGCA   3180
GGAGGAAATCTCTATCTCTGGCTGCAAGATGAGGCTGAGCTACCTGAGCAGCCGGACCCC   3240
TGGCTACAAATCTGTCCTGAGGATCAGCCTCACCCACCCGACCATCCCCTTCAACCTCAT   3300
GAAGGTGCACCTCATGGTAGCGGTGGAGGGCCGCCTCTTCAGGAAGTGGTTCGCTGCAGC   3360
CCCAGACCTGTCCTATTATTTCATTTGGGACAAGACAGACGTCTACAACCAGAAGGTGTT   3420
TGGGCTTTCAGAAGCCTTTGTTTCCGTGGGTTATGAATATGAATCCTGCCCAGATCTAAT   3480
CCTGTGGGAAAAAAGAACAACAGTGCTGCAGGGCTATGAAATTGACGCGTCCAAGCTTGG   3540
AGGATGGAGCCTAGACAAACATCATGCCCTCAACATTCAAAGTGGTATCCTGCACAAAGG   3600
GAATGGGGAGAACCAGTTTGTGTCTCAGCAGCCTCCTGTCATTGGGAGCATCATGGGCAA   3660
TGGGCGCCGGAGAAGCATCTCCTGCCCCAGCTGCAACGGCCTTGCTGACGGCAACAAGCT   3720
CCTGGCCCCAGTGGCCCTCACCTGTGGCTCTGACGGGAGCCTCTATGTGGGTGATTTCAA   3780
CTACATTAGAAGGATCTTCCCCTCTGGAAATGTCACCAACATCCTAGAGCTGAGGAATAA   3840
AGATTTCAGACATAGTCACAGTCCAGCACACAAATACTACCTGGCCACAGACCCCATGAG   3900
TGGGGCCGTCTTCCTTTCTGACAGCAACAGCCGGCGGGTCTTTAAAATCAAGTCCACTGT   3960
GGTGGTGAAGGACCTTGTCAAGAACTCTGAGGTGGTTGCGGGGACAGGTGACCAGTGCCT   4020
CCCCTTTGATGACACTCGCTGCGGGGATGGTGGGAAGGCCACAGAAGCCACACTCACCAA   4080
TCCCAGGGGTATTACAGTGGACAAGTTTGGGCTGATCTACTTCGTGGATGGCACCATGAT   4140
CAGACGCATCGATCAGAATGGGATCATCTCCACCCTGCTCGGCTCTAATGATCTCACATC   4200
AGCCCGGCCACTCAGCTGTGATTCTGTCATGGATATTTCCCAGGTAAGACTGGAGTGGCC   4260
CACAGACTTAGCCATCAACCCAATGGACAACTCACTTTATGTCCTCGACAACAATGTGGT   4320
CCTGCAAATCTCTGAAAACCACCAGGTGCGCATTGTCGCCGGGAGGCCCATGCACTGCCA   4380
GGTCCCTGGCATTGACCACTTCCTGCTAAGCAAGGTGGCCATCCACGCAACCCTGGAGTC   4440
AGCCACCGCTTTGGCTGTTTCACACAATGGGGTCCTGTATATTGCTGAGACTGATGAGAA   4500
AAAGATCAACCGCATCAGGCAGGTCACCACTAGTGGAGAGATCTCACTCGTTGCTGGGGC   4560
CCCCAGTGGCTGTGACTGTAAAAATGATGCCAACTGTGATTGTTTTTCTGGAGACGATGG   4620
TTATGCCAAGGATGCAAAGTTAAATACCCCATCTTCCTTGGCTGTGTGTGCTGATGGGGA   4680
GCTCTACGTGGCCGACCTTGGGAACATCCGAATTCGGTTTATCCGGAAGAACAAGCCTTT   4740
CCTCAACACCCAGAACATGTATGAGCTGTCTTCACCAATTGACCAGGAGCTCTATCTGTT   4800
TGATACCACCGGCAAGCACCTGTACACCCAAAGCCTGCCCACAGGAGACTACCTGTACAA   4860
CTTCACCTACACTGGGGACGGCGACATCACACTCATCACAGACAACAATGGCAACATGGT   4920
AAATGTCCGCCGAGACTCTACTGGGATGCCCCTCTGGCTGGTGGTCCCAGATGGCCAGGT   4980
GTACTGGGTGACCATGGGCACCAACAGTGCACTCAAGAGTGTGACCACACAAGGACACGA   5040
GTTGGCCATGATGACATACCATGGCAATTCCGGCCTTCTGGCAACCAAAAGCAATGAAAA   5100
CGGATGGACAACATTTTATGAGTACGACAGCTTTGGCCGCCTGACAAATGTGACCTTCCC   5160
TACTGGCCAGGTGAGCAGTTTCCGAAGTGATACAGACAGTTCAGTGCATGTCCAGGTAGA   5220
GACCTCCAGCAAGGATGATGTCACCATAACCACCAACCTGTCTGCCTCAGGCGCCTTCTA   5280
CACACTGCTGCAAGACCAAGTCCGGAACAGCTACTACATCGGGGCCGATGGCTCCTTGCG   5340
GCTGCTGCTGGCCAACGGCATGGAGGTGGCGCTGCAGACTGAGCCCCACTTGCTGGCTGG   5400
CACCGTCAACCCCACCGTGGGCAAGAGGAATGTCACGCTGCCCATCGACAACGGCCTCAA   5460
CCTGGTGGAGTGGCGCCAGCGCAAAGAGCAGGCTCGGGGCCAGGTCACTGTCTTTGGGCG   5520
CCGGCTGCGGGTGCACAACCGAAATCTCCTATCTCTGGACTTTGATCGCGTAACACGCAC   5580
AGAGAAGATCTATGATGACCACCGCAAGTTCACCCTTCGGATTCTGTACGACCAGGCGGG   5640
GCGGCCCAGCCTCTGGTCACCCAGCAGCAGGCTGAATGGTGTCAACGTGACATACTCCCC   5700
TGGGGGTTACATTGCTGGCATCCAGAGGGGCATCATGTCTGAAAGAATGGAATACGACCA   5760
GGCGGGCCGCATCACATCCAGGATCTTCGCTGATGGGAAGACATGGAGCTACACATACTT   5820
AGAGAAGTCCATGGTGCTGCTACTACACAGCCAGAGGCAGTATATCTTTGAGTTCGACAA   5880
GAATGACCGCCTCTCTTCTGTGACGATGCCCAACGTGGCGCGGCAGACACTAGAGACCAT   5940
CCGCTCAGTGGGCTACTACAGAAACATCTATCAGCCCCCTGAGGGCAATGCCTCAGTCAT   6000
ACAGGACTTCACTGAGGATGGGCACCTCCTTCACACCTTCTACCTGGGCACTGGCCGCAG   6060
GGTGATATACAAGTATGGCAAACTGTCAAAGCTGGCAGAGACGCTCTATGACACCACCAA   6120
GGTCAGTTTCACCTATGACGAGACGGCAGGCATGCTGAAGACCATCAACCTACAGAATGA   6180
GGGCTTCACCTGCACCATCCGCTACCGTCAGATTGGGCCCCTGATTGACCGACAGATCTT   6240
CCGCTTCACTGAGGAAGGCATGGTCAACGCCCGTTTTGACTACAACTATGACAACAGCTT   6300
```

```
CCGGGTGACCAGCATGCAGGCTGTGATCAACGAGACCCCACTGCCCATTGATCTCTATCG   6360
CTATGATGATGTGTCAGGCAAGACAGAGCAGTTTGGGAAGTTTGGTGTCATTTACTATGA   6420
CATTAACCAGATCATCACCACAGCTGTCATGACCCACACCAAGCATTTTGATGCATATGG   6480
CAGGATGAAGGAAGTGCAGTATGAGATCTTCCGCTCGCTCATGTACTGGATGACCGTCCA   6540
GTATGATAACATGGGGCGAGTAGTGAAGAAGGAGCTGAAGGTAGGACCCTACGCCAATAC   6600
CACTCGCTACTCCTATGAGTATGATGCTGACGGCCAGCTGCAGACAGTCTCCATCAATGA   6660
CAAGCCACTCTGGCGCTACAGCTACGACCTCAATGGGAACCTGCACTTACTGAGCCCTGG   6720
GAACAGTGCACGGCTCACACCACTACGGTATGACATCCGCGACCGCATCACTCGGCTGGG   6780
TGACGTGCAATACAAGATGGATGAGGATGGCTTCCTGAGGCAGCGGGGCGGTGATATCTT   6840
TGAGTACAACTCAGCTGGCCTGCTCATCAAGGCCTACAACCGGGCTGGCAGCTGGAGTGT   6900
CAGGTACCGCTACGATGGCCTGGGGCGGCGCGTGTCCAGCAAGAGCAGCCACAGCCACCA   6960
CCTGCAGTTCTTCTATGCAGACCTGACCAACCCCACCAAGGTCACCCACCTGTACAACCA   7020
CTCCAGCTCTGAGATCACCTCCCTCTACTACGACTTGCAAGGACACCTCTTTGCCATGGA   7080
GCTGAGCAGTGGTGATGAGTTTTACATAGCTTGTGACAACATCGGGACCCCTCTTGCTGT   7140
CTTTAGTGGAACAGGTTTGATGATCAAGCAAATCCTGTACACAGCCTATGGGGAGATCTA   7200
CATGGATACCAACCCCAACTTTCAGATCATCATAGGCTACCATGGTGGCCTCTATGATCC   7260
ACTCACCAAGCTTGTCCACATGGGCCGGCGAGATTATGATGTGCTGGCCGGACGCTGGAC   7320
TAGCCCAGACCACGAGCTGTGGAAGCACCTTAGTAGCAGCAACGTCATGCCTTTTAATCT   7380
CTATATGTTCAAAAACAACAACCCCATCAGCAACTCCCAGGACATCAAGTGCTTCATGAC   7440
AGATGTTAACAGCTGGCTGCTCACCTTTGGATTCCAGCTACACAACGTGATCCCTGGTTA   7500
TCCCAAACCAGACATGGATGCCATGGAACCCTCCTACGAGCTCATCCACACACAGATGAA   7560
AACGCAGGAGTGGGACAACAGCAAGTCTATCCTCGGGGTACAGTGTGAAGTACAGAAGCA   7620
GCTCAAGGCCTTTGTCACCTTAGAACGGTTTGACCAGCTCTATGGCTCCACAATCACCAG   7680
CTGCCAGCAGGCTCCAAAGACCAAGAAGTTTGCATCCAGCGGCTCAGTCTTTGGCAAGGG   7740
GGTCAAGTTTGCCTTGAAGGATGGCCGAGTGACCACAGACATCATCAGTGTGGCCAATGA   7800
GGATGGGCGAAGGGTTGCTGCCATCTTGAACCATGCCCACTACCTAGAGAACCTGCACTT   7860
CACCATTGATGGGGTGGATACCCATTACTTTGTGAAACCAGGACCTTCAGAAGGTGACCT   7920
GGCCATCCTGGGCCTCAGTGGGGGGCGGCGAACCCTGGAGAATGGGGTCAACGTCACTGT   7980
GTCCCAGATCAACACAGTACTTAATGGCAGGACTAGACGCTACACAGACATCCAGCTCCA   8040
GTACGGGGCACTGTGCTTGAACACACGCTACGGGACAACGTTGGATGAGGAGAAGGCACG   8100
GGTCCTGGAGCTGGCCCGGCAGAGAGCCGTGCGCCAAGCGTGGGCCCGCGAGCAGCAGAG   8160
ACTGCGGGAAGGGGAGGAAGGCCTGCGGGCCTGGACAGAGGGGGAGAAGCAGCAGGTGCT   8220
GAGCACAGGGCGGGTGCAAGGCTACGACGGCTTTTTCGTGATCTCTGTCGAGCAGTACCC   8280
AGAACTGTCAGACAGCGCCAACAACATCCACTTCATGAGACAGAGCGAGATGGGCCGGAG   8340
GTGACAGAGAGGAC
```

A disclosed NOV4 nucleic acid maps to chromosome 11, and is found in at least brain, spinal chord, testis, heart, lung, parathyroid, stomach, breast, colon, epidermis, ovary and kidney. A NOV4 nucleic acid has 7504 of 8359 bases (89%) identical to a gb:GENBANK-ID:AB025413|acc: AB025413.1 mRNA from Mus musculus TEN-M4.

A NOV4 polypeptide (SEQ ID NO:14) encoded by SEQ ID NO:13 is 2769 amino acid residues and is presented using the one letter code in Table 4B. Signal P, Psort and/or Hydropathy results predict that NOV4 does not have a signal peptide and is likely to be localized mitochondrial inner membrane with a certainty of 0.8363. In other embodiments, NOV4 may also be localized to the plasma membrane with a certainty of 0.65 or to the nucleus with a certainty of 0.6000, or microbody with a certainty of 0.3936.

Table 4B. NOV4 Polypeptide SEQ ID NO:14	
MDVKERKPYRSLTRRRDAERRYTSSSADSEEGKAPQKSYSSSETLKAYDQDARLAYGSRV	60
KDIVPQEAEEFCRTGANFTLRELGLEEVTPPHGTLYRTDIGLPQCGYSMGAGSDADMEAD	120
TVLSPEHPVRLWGRSTRSGRSSCLSSRANSNLTLTDTEHENTETDHPGGLQNHARLRTPP	180
PPLSHAHTPNQHHAASINSLNRGNFTPRSNPSPAPTDHSLSGEPPAGGAQEPAHAQENWL	240
LNSNIPLETRNLGKQPFLGTLQDNLIEMDILGASRHDGAYSDGHFLFKPGGTSPLFCTTS	300
PGYPLTSSTVYSPPPRPLPRSTFARPAFNLKKPSKYCNWKCAALSAIVISATLVILLAYF	360
VAMHLFGLNWHLQPMEGQMYEITEDTASSWPVPTDVSLYPSGGTGLETPDRKGKGTTEGK	420
PSSFFPEDSFIDSGEIDVGRRASQKIPPGTFWRSQVFIDHPVHLKFNVSLGKAALVGIYG	480
RKGLPPSHTQFDFVELLDGRRLLTQEARSLEGTPRQSRGTVPPSSHETGFIQYLDSGIWH	540
LAFYNDGKESEVVSFLTTAIESVDNCPSNCYGNGDCISGTCHCFLGFLGPDCGRASCPVL	600
CSGNGQYMKGRCLCHSGWKGAECDVPTNQCIDVACSNHGTCITGTCICNPGYKGESCEEV	660
DCMDPTCSGRGVCVRGECHCFVGWGGTNCETPRATCLDQCSGHGTFLPDTGLCSCDPSWT	720
GHDCSIEICAADCGGHGVCVGGTCRCEDGWMGAACDQRACHPRCAEHGTCRDGKCECSPG	780
WNGEHCTIAHYLDRVVKEGCPGLCNGNGRCTLDLNGWHCVCQLGWRGAGCDTSMETACGD	840
SKDNDGDGLVDCMDPDCCLQPLCHINPLCLGSPNPLDIIQETQVPVSQQNLHSFYDRIKF	900
LVGRDSTHIIPGENPFDGGHACVIRGQVMTSDGTPLVGVNISFVNNPLFGYTISRQDGSF	960
DLVTNGGISIILRFERAPFITQEHTLWLPWDRFFVMETIIMRHEENEIPSCDLSNFARPN	1020
PVVSPSPLTSFASSCAEKGPIVPEIQALQEEISISGCKMRLSYLSSRTPGYKSVLRISLT	1080
HPTIPFNLMKVHLMVAVEGRLFRKWFAAAPDLSYYFIWDKTDVYNQKVFGLSEAFVSVGY	1140
EYESCPDLILWEKRTTVLQGYEIDASKLGGWSLDKHHALNIQSGILHKGNGENQFVSQQP	1200
PVIGSIMGNGRRRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDFNYIRRIFPSGNV	1260
TNILELRNKDFRHSHSPAHKYYLATDPMSGAVFLSDSNSRRVFKIKSTVVVKDLVKNSEV	1320
VAGTGDQCLPFDDTRCGDGGKATEATLTNPRGITVDKFGLIYFVDGTMIRRIDQNGIIST	1380
LLGSNDLTSARPLSCDSVMDISQVRLEWPTDLAINPMDNSLYVLDNNVVLQISENHQVRI	1440
VAGRPMHCQVPGIDHFLLSKVAIHATLESATALAVSHNGVLYIAETDEKKINRIRQVTTS	1500
GEISLVAGAPSGCDCKNDANCDCFSGDDGYAKDAKLNTPSSLAVCADGELYVADLGNIRI	1560
RFIRKNKPFLNTQNMYELSSPIDQELYLFDTTGKHLYTQSLPTGDYLYNFTYTGDGDITL	1620
ITDNNGNMVNVRRDSTGMPLWLVVPDGQVYWVTMGTNSALKSVTTQGHELAMMTYHGNSG	1680
LLATKSNENGWTTFYEYDSFGRLTNVTFPTGQVSSFRSDTDSSVHVQVETSSKDDVTITT	1740
NLSASGAFYTLLQDQVRNSYYIGADGSLRLLLANGMEVALQTEPHLLAGTVNPTVGKRNV	1800
TLPIDNGLNLVEWRQRKEQARGQVTVFGRRLRVHNRNLLSLDFDRVTRTEKIYDDHRKFT	1860
LRILYDQAGRPSLWSPSSRLNGVNVTYSPGGYIAGIQRGIMSERMEYDQAGRITSRIFAD	1920
GKTWSYTYLEKSMVLLLHSQRQYIFEFDKNDRLSSVTMPNVARQTLETIRSVGYYRNIYQ	1980
PPEGNASVIQDFTEDGHLLHTFYLGTGRRVIYKYGKLSKLAETLYDTTKVSFTYDETAGM	2040
LKTINLQNEGFTCTIRYRQIGPLIDRQIFRFTEEGMVNARFDYNYDNSFRVTSMQAVINE	2100
TPLPIDLYRYDDVSGKTEQFGKFGVIYYDINQIITTAVMTHTKHFDAYGRMKEVQYEIFR	2160
SLMYWMTVQYDNMGRVVKKELKVGPYANTTRYSYEYDADGQLQTVSINDKPLWRYSYDLN	2220
GNLHLLSPGNSARLTPLRYDIRDRITRLGDVQYKMDEDGFLRQRGGDIFEYNSAGLLIKA	2280
YNRAGSWSVRYRYDGLGRRVSSKSSHSHHLQFFYADLTNPTKVTHLYNHSSSEITSLYYD	2340
LQGHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIKQILYTAYGEIYMDTNPNFQIII	2400
GYHGGLYDPLTKLVHMGRRDYDVLAGRWTSPDHELWKHLSSSNVMPFNLYMFKNNNPISN	2460
SQDIKCFMTDVNSWLLTFGFQLHNVIPGYPKPDMDAMEPSYELIHTQMKTQEWDNSKSIL	2520
GVQCEVQKQLKAFVTLERFDQLYGSTITSCQQAPKTKKFASSGSVFGKGVKFALKDGRVT	2580
TDIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFVKPGPSEGDLAILGLSGGRRT	2640
LENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCLNTRYGTTLDEEKARVLELARQRAVR	2700
QAWAREQQRLREGEEGLRAWTEGEKQQVLSTGRVQGYDGFFVISVEQYPELSDSANNIHF	2760
MRQSEMGRR	

The full amino acid sequence of the protein of the invention was found to have 2688 of 2771 amino acid residues (97%) identical to, and 2728 of 2771 amino acid residues (98%) similar to, the 2771 amino acid residue ptnr:SPTREMBL-ACC:Q9WTS7 protein from Mus musculus TEN-M4.

NOV4 also has homology to the amino acid sequences shown in the BLASTP data listed in Table 4C.

Table 4C. BLAST results for NOV4

Gene Index/ Identifier	Protein/ Organism	Length (aa)	Identity (%)	Positives (%)	Expect
gi|16551957|dbj|BAB 71206.1| (AK056531)	unnamed protein product [Homo sapiens]	730	99	99	0.0
gi|7657417|ref|NP 035987.2| (NM_011857)	odd Oz/ten-m homolog 3 (Drosophila); odd Oz/ten-m homolog 1 (Drosophila) [Mus musculus]	2715	66	79	0.0
gi|13649010|ref|X P_010128.3| XM_010128	odz (odd Oz/ten-m, Drosophila) homolog 1 [Homo sapiens]	2725	62	76	0.0
gi|1079143|pir||S 47008	tenascin-like protein - fruit fly (Drosophila melanogaster)	2515	33	53	0.0
gi|8922444|ref|NP 060574.1| (NM_018104)	hypothetical protein FLJ10474; hypothetical protein FLJ10886 [Homo sapiens]	1045	99	99	0.0

The homology of these sequences is shown graphically in the ClustalW analysis shown in Table 4D.

Table 4D ClustalW Analysis of NOV4

Tables 4E lists the domain description from DOMAIN analysis results against NOV4. This indicates that the NOV4 sequence has properties similar to those of other proteins known to contain this domain.

```
1) NOV4 (SEQ ID NO:13)
2) gi|16551957 (SEQ ID NO:50)
3) gi|7657417 (SEQ ID NO:51)
4) gi|13649010 (SEQ ID NO:52)
5) gi|1079143 (SEQ ID NO:53)
6) gi|8922444 (SEQ ID NO:54)
10         20          30          40          50
                  ....|....|....|....|....|....|....|....|....|....|
NOV4              MDVKERKPYRSLT-RRRDAERRYTSSSADSEEGKAP-QKSYSSSETLKAY
gi|16551957|      --------------------------------------------------
gi|7657417|       MDVKERRPYCSLTKSRREKERRYTNSSADNEECRVPTQKSYSSSETLKAF
gi|13649010|      MEQTDCKPYQPLPKVKHEMDLAYTSSSDESEDGRKP-RQSYNSRETLHEY
```

```
gi|1079143|    --------------------------------------------------
gi|8922444|    --------------------------------------------------

                        60        70        80        90       100
               ....|....|....|....|....|....|....|....|....|....|
NOV4           DQD-ARLAYGSRVKDIVPQEAEEFCRTGANFTLRECGLECVTPPHGTLYR
gi|16551957|   --------------------------------------------------
gi|7657417|    DHDYSRLLYGNRVKDLVHREADEYTRQGQNFTLRQCGVCCSATRRGVAFC
gi|13649010|   NQELR----------MN-YNSQSRKRKEVEKSTQEMEFCRTSHTLCSGYQ
gi|1079143|    -----------------------------------CNFRCDLVARCSSPW
gi|8922444|    --------------------------------------------------

                       110       120       130       140       150
               ....|....|....|....|....|....|....|....|....|....|
NOV4           TDCG-CPQCGYSCGAGSDADMEADTVLSPEHPVRCWGRSTRSGRSSCLSS
gi|16551957|   --------------------------------------------------
gi|7657417|    AECG-CPHRGYSCSAGSDADTENEAVMSPEHAMRCWGRGVKSGRSSCLSS
gi|13649010|   TDCHSCSRHGYQCEMGSDVDTETEGAASPDHALRCWIRGMKSEHSSCLSS
gi|1079143|    PGCGSCSVLFAFCVMLILLTTTGVIKWCQSPPCSCLVGNEASEVSSAKSS
gi|8922444|    --------------------------------------------------

                       160       170       180       190       200
               ....|....|....|....|....|....|....|....|....|....|
NOV4           RACSNLTLTDCEHENTETDH------------------------PGGLQN
gi|16551957|   --------------------------------------------------
gi|7657417|    RSCSALTLTDCEHENRSDSE------------------------SEQPSN
gi|13649010|   RACSALSLTDCDHERKSDGENGFKFSPVCCDMEAQAGSTQDVQSSPHNQF
gi|1079143|    NTCLSKLHNSCVRAKNGQGIG-------------------------LAQG
gi|8922444|    --------------------------------------------------

                       210       220       230       240       250
               ....|....|....|....|....|....|....|....|....|....|
NOV4           HARLRTPPPPLSHAHTPNQHHAASINSLNRGNFTPRSNPSPSPTDHSLSG
gi|16551957|   --------------------------------------------------
gi|7657417|    NPGQPTLQPLPPSHKQHPAQHHPSITSLNRNSLTNRRNQSPSPPAALPSE
gi|13649010|   TFRPLPPPPPPPHACTCARKPPPAADSLQRRSMTTRSQPSPSAPAPPTST
gi|1079143|    QSGLGAAGVGSGGGSSAATVTTATSNSGTAQGLQSTSASAESTSSAATSS
gi|8922444|    --------------------------------------------------

                       260       270       280       290       300
               ....|....|....|....|....|....|....|....|....|....|
NOV4           EPPAGGAQEPAHAQENWLLNSNIPLETRNLGKQPFLGTLQDNLIEMDILG
gi|16551957|   --------------------------------------------------
gi|7657417|    LQTTP---ESVQLQDSWVLGSNVPLESR----------------------
gi|13649010|   QDS-------VHLHNSWVLNSNIPLETR----------------------
gi|1079143|    SQS-----------------------------------------------
gi|8922444|    --------------------------------------------------

                       310       320       330       340       350
               ....|....|....|....|....|....|....|....|....|....|
NOV4           ASRHDGAYSDGHFLFKPGGTSPLFCTTSPGYPLTSSTVYSPPPRPLERSS
gi|16551957|   --------------------------------------------------
```

```
gi|7657417|    ----------HFLFKTGTGTTPLFSTALPGYTMASGSVYSPPTRPLPRNT
gi|13649010|   ----------HFLFKHGSGSSAIFSAASQNYPLTSNTVYSPPPRPLPRST
gi|1079143|    ------------------------S-LLPSLSSSLANANNGGARTFPARS
gi|8922444|    --------------------------------------------------

                       360       370       380       390       400
               ....|....|....|....|....|....|....|....|....|....|
NOV4           FARPAFNLKKPSKYCNWKCAALSAIVISATLVILLAYFVAMHLFGLNWHL
gi|16551957|   --------------------------------------------------
gi|7657417|    LSRSAFKFKKSSKYCSWRCTALCAVGVSVLLAILLSYFIAMHLFGLNWHL
gi|13649010|   FSRPAFTFNKPYRCCNWKCTALSATAITVTLALLLAYVIAVHLFGLTWQL
gi|1079143|    FP--------------------------------------P---------
gi|8922444|    --------------------------------------------------

                       410       420       430       440       450
               ....|....|....|....|....|....|....|....|....|....|
NOV4           QPMEGQMYEITEDTASSWPVPTDVSLYPSGGTGLETPDRKGKGTTEGKPS
gi|16551957|   --------------------------------------------------
gi|7657417|    QQTENDTFENGKVNSD--TVPTNTVSLPSGDN---------GKLG-----
gi|13649010|   QPVEGELYANGVSKGNRGTESMDTTYSPIGGKVS------DKSEK-----
gi|1079143|    ----------------------------DGTTFG----------------
gi|8922444|    --------------------------------------------------

                       460       470       480       490       500
               ....|....|....|....|....|....|....|....|....|....|
NOV4           SFFPEDSFIDSGELDVGRRASDKLPPGTFFRSGVFIDHFVHLKFNVSEGK
gi|16551957|   --------------------------------------------------
gi|7657417|    GFTHENNTIDSGKLDICRRAIDEVPPGIFFRSGLFIDQFQFLNFNISLQK
gi|13649010|   KVFQKGRAIDTGEFDICARVMGTIPPGLFFRFGITIHHFIYLKFNISLAK
gi|1079143|    ------------QSTLGQKLTKEIQPYSVFNMGFYQSEFAYVKFDYSIPG
gi|8922444|    --------------------------------------------------

                       510       520       530       540       550
               ....|....|....|....|....|....|....|....|....|....|
NOV4           ARLVGIYGRKGLPPSHTCFDFVELLDGRRLLTQEARSLEGTPRQSRGTVP
gi|16551957|   --------------------------------------------------
gi|7657417|    DALLGVYGRKGLPPSHTCYDFVELLDGSRLIAREQRNLVEHERAGRQARS
gi|13649010|   DSLLGIYGRRNIPPIHTCFDFVKLNDGKQLVKQDSKGSDDTQHSPRNLIL
gi|1079143|    GASLGVYGRRNALPTHTCYHKKEVLSGFSASTRTARAAHLSITR------
gi|8922444|    --------------------------------------------------

                       560       570       580       590       600
               ....|....|....|....|....|....|....|....|....|....|
NOV4           PSSHETGFIQYLDSGINHLAFYND-GKESGVVSFLTTAIESVDNCPSNCY
gi|16551957|   --------------------------------------------------
gi|7657417|    VSLHEAGFIQYLDSGINHLAFYND-GKNPGQVSFNTIVIESVVECPRNCH
gi|13649010|   TSLQETGFIEEMDQGPNYLAFYND-GKKMGQVFVLTTAIEIMDDCSTNCN
gi|1079143|    ------EVTRSMKPGHNFVSLYNDDGDVQGLTFYAAVAEDMTQNCPNGCS
gi|8922444|    --------------------------------------------------

                       610       620       630       640       650
               ....|....|....|....|....|....|....|....|....|....|
```

```
NOV4              GNGDCISGTCHCFLGELGPDCGRASCPVLCSGNGQYMKGRCLCHSGWKGA
gi|16551957|      --------------------------------------------------
gi|7657417|       GNGECVSGTCHCFPGELGPDCSRAACPVLCSGNGQYSKGRCLCFSGWKGT
gi|13649010|      GNGECLSGHCHCFPGELGPDCARDSCPVLCGGNGEYEKGHCVCRHGWKGP
gi|1079143|       GNGQCLLGHCQCNPGFGGDDCSESVCPVECSQHGEYTNGECICNPGWKGK
gi|8922444|       --------------------------------------------------

                          660       670       680       690       700
                  ....|....|....|....|....|....|....|....|....|....|
NOV4              ECDVPTNQCIDVACSNHGTCITGTCICNPGYKGESCEEVECMDPTCSGRG
gi|16551957|      --------------------------------------------------
gi|7657417|       ECDVPTTQCIDPQCGGRGICIMGSCACNSGYKGENCEEADCLDPGCSNHG
gi|13649010|      ECDVPEEQCIDPTCFGHGTCIMGVCICVPGYKGRICEEEDCLDPMCSNHG
gi|1079143|       ECSLRHDECKVADCSGHGHCVSGKCQCMRGYKGKFCEEVECPHPNCSGHG
gi|8922444|       --------------------------------------------------

                          710       720       730       740       750
                  ....|....|....|....|....|....|....|....|....|....|
NOV4              VCVRGECHCFVGNGGTNCETP---RATCLDQCSGHGTFLPETGLCSCEPS
gi|16551957|      --------------------------------------------------
gi|7657417|       VCIHGECHCNPGYGGSNCEIL---KTMCADQCSGHGTYLQESGSCTCDPN
gi|13649010|      ICVKGECHCSTGYGGVNCETP---LPVCQEQCSGHGTFLLQAGVCSCEPK
gi|1079143|       FCADGTCICKKGYKGPDCATMDQDALQCLPDCSGHGTFDLQTQTCTCEAK
gi|8922444|       --------------------------------------------------

                          760       770       780       790       800
                  ....|....|....|....|....|....|....|....|....|....|
NOV4              NTGHDCSIEYCAADCGGHGVCVGGTCRCEDGNMCAACDQRACHPRCAEHG
gi|16551957|      --------------------------------------------------
gi|7657417|       NTGPDCSNETCSVDCGSHGVCMGGSCRCEEGNTCPACNQRACHPRCAEHG
gi|13649010|      NTGSDCSTEKCTMECGSHGVCSRGICQCEEGNVCPTCEERSCHSHCIEHG
gi|1079143|       NSGDDCSKERCDLDCGQHGRCEGDACACDPENGCEYCNTRLCDVRCNEHG
gi|8922444|       --------------------------------------------------

                          810       820       830       840       850
                  ....|....|....|....|....|....|....|....|....|....|
NOV4              TCRDGKCECSPGWNGEHCTIAHYLDRVVKEGCPGLCNGNGRCTLDLNG-W
gi|16551957|      --------------------------------------------------
gi|7657417|       TCKDGKCECSQGWNEKHCTIAHYLDKIVKEGCPGLCNSNERCTLDQNG-W
gi|13649010|      QCKDGKCECSPGRNGDHCTIAHYLD-AVRDGCPGLCFGNGRCTLDQNG-W
gi|1079143|       QCKNGTCLCVTGWNGKHCTEEG---------CENSCAGHGQCRVSGEGQW
gi|8922444|       --------------------------------------------------

                          860       870       880       890     . 900
                  ....|....|....|....|....|....|....|....|....|....|
NOV4              HCVCQLGWRGAGCDTSMETACGDSKDNLGDGEVDCMDPDCCLQPLCHINP
gi|16551957|      --------------------------------------------------
gi|7657417|       HCVCQPGWRGAGCDVAMETLCTDSKDNEGDGEIDCMDPDCCLQSSCQNQP
gi|13649010|      HCVCQVGWSGTGCNVVMEMLCGDNLDNLGDGETDCVDPDCCQQSNCYISP
gi|1079143|       ECRCYEGWDGPDCGIALELNCGDSKDNDKDGEVDCEDPECCASHVCKTSQ
gi|8922444|       --------------------------------------------------
```

```
                        910       920       930       940       950
                ....|....|....|....|....|....|....|....|....|....|
NOV4            LCLGSPNELDIIQETQVPVSQQNLHSFYDRIKFEVGRDSTHIIPGENPED
gi|16551957|    --------------------------------------------------
gi|7657417|     YCRGLPDPQDIISQSLQTPSQQAAKSFYDRISFLIGSDSTHVLPGESPEN
gi|13649010|    LCQGSPDPLDLIQQSQTLFSQHTSRLFYDRIKFLIGKDSTHVIPPEVSED
gi|1079143|     LCVSAPKPIDVILRKQP---PAITASFFERMKFLIDESSLQNYAKLETEN
gi|8922444|     --------------------------------------------------

                        960       970       980       990      1000
                ....|....|....|....|....|....|....|....|....|....|
NOV4            GGHRCVIRGQVMTSDGTPLVGVNISFVNNPLFGYTISRQDGSFDLVTNGG
gi|16551957|    --------------------------------------------------
gi|7657417|     KSLASVIRGQVLTADGTPLIGVNVSFLHYSEYGYTLTRQDGMFDLVANGG
gi|13649010|    SRRACVIRGQVVAIDGTPLVGVNVSFLHHSDYGFTISRQDGSFDLVAIGG
gi|1079143|     ESRKAVIRGRVVTSLGMGLVGVRVSTTTLLEG-FTLTRDDGWFDLMVNGG
gi|8922444|     --------------------------------------------------

                       1010      1020      1030      1040      1050
                ....|....|....|....|....|....|....|....|....|....|
NOV4            ISAILRFERAPFITQEHTLWLPWQRFFVMETEIMRH------EENKIPSC
gi|16551957|    --------------------------------------------------
gi|7657417|     ASETLVFERSPFLTQYHTVWLPWNVFYVMETLVMKK------EENDIPSC
gi|13649010|    ISVILIFDESPFLPRKRTLWLPWNQFIVVSKVTMQR------VVSDPPSC
gi|1079143|     GAVTLQFGRAPFRPQSRIVQVPWNEVVIKDLVVMSMSEEKGLAVTTTHTC
gi|8922444|     --------------------------------------------------

                       1060      1070      1080      1090      1100
                ....|....|....|....|....|....|....|....|....|....|
NOV4            DLSNEARPNPVVSPSPLTSGASSCAEKGPIVPEISALQEEISESGCKWRE
gi|16551957|    --------------------------------------------------
gi|7657417|     DLSGEVRPSPIIVSSPLSTGFRSSPEDSPEIPETQVIHEETTIPGTDRKE
gi|13649010|    DISNEISPNPIVLPSPLTSEGGSCPERGTEVPELCVVQEEIPEPSSFPRE
gi|1079143|     FAHDVDLMKPVVLASWKHGEQGACPGRSAIIAESCVIQESLQEPGTGLNE
gi|8922444|     --------------------------------------------------

                       1110      1120      1130      1140      1150
                ....|....|....|....|....|....|....|....|....|....|
NOV4            SYLSSRTPGYKSVLRESLEHPTEEFNLMKVHEMVAVEERIFPRKWEAEAPD
gi|16551957|    --------------------------------------------------
gi|7657417|     SYLSSRAAGYKSVLKETMIQAVEEFNLMKVHLMVAVVERIFQRNFPESPN
gi|13649010|    SYLSSRTPGYKTLLRYLLEHSTEEVGMIKVHLTVAMEERETQKWEPSAIN
gi|1079143|     VYHSSRAAGYLSTIKLQLEPDVEETSLHLIHERITYEESIEFERIEEDPG
gi|8922444|     --------------------------------------------------

                       1160      1170      1180      1190      1200
                ....|....|....|....|....|....|....|....|....|....|
NOV4            ISYYFIRDKTDVYNGKVYGLSEAFVSVGYEYESCPELILEEKRITVLQSY
gi|16551957|    --------------------------------------------------
gi|7657417|     HAYTEIRDKTDAYNGKVYGLSEAVVSVGYEYESCLDLTLEEKRFAVLQSY
gi|13649010|    LVSTEARNKTDIIGGKVVCLAEALVSVGYEYETCPDFILKEQRFVVLQSF
gi|1079143|     IKETYARNRLAIYRGRVYGVTTEVVKVGYQYTDCTD-IVNIIQITKLSCH
```

```
gi|8922444|      --------------------------------------------------

                          1210      1220      1230      1240      1250
                 ....|....|....|....|....|....|....|....|....|....|
NOV4             EIDASKLGGNSLDKHHALNIQSEIIHKGNEKNQFVSQQEPVIGSIMGNER
gi|16551957|     --------------------------------------------------
gi|7657417|      ELDASNMGGNIEDKHHVLLVQNSILYKGNEENQFISQQEPVVSSIMGNER
gi|13649010|     ENDASNLGGHSLNKHHILNPQSEIIHKGNEENMPISQQEPVISSIMGNEH
gi|1079143|      DMSISEVGGWNELIHHRYNFHEEILQKGDESNIYIRNKERTILITMGDEH
gi|8922444|      --------------------------------------------------

                          1260      1270      1280      1290      1300
                 ....|....|....|....|....|....|....|....|....|....|
NOV4             RRSISCPSCNGLADGNKLLAPVALTCGSDGSLYVGDFNYIRRIFPSGNVT
gi|16551957|     --------------------------------------------------
gi|7657417|      KRSISCPSCNGQADGNKLLAPVALACGIDGSLYVGDFNYVRRIFPSGNVT
gi|13649010|     DRSVACTNCNGPAHNNKLFAPVALASGPDGSVYVGDFNFVRRIFPSGNSV
gi|1079143|      DRPLECPDCDGQATKQRLLAPVALAAAPDGSLFVGDFNYIRRIMTDGSIR
gi|8922444|      --------------------------------------------------

                          1310      1320      1330      1340      1350
                 ....|....|....|....|....|....|....|....|....|....|
NOV4             NELSLRNKDFRHSHSPAHKYYLETDEMSGALFLSDSNGRRVFRIISTVVV
gi|16551957|     --------------------------------------------------
gi|7657417|      SVLSLRNKDFRHSSNPAHRYYLETDEVTGDLYVSDINGRRIYHPKSLTGA
gi|13649010|     SIQELS-------TSPAHKEYLEMDEVSESLYLSDINTRKVYELSSLVET
gi|1079143|      TEVKLN------ATRVSYRYHMELSELDGTLYVSDPESHQEIRVEDTNDY
gi|8922444|      --------------------------------------------------

                          1360      1370      1380      1390      1400
                 ....|....|....|....|....|....|....|....|....|....|
NOV4             KDLVKNSEVVAGTGDQCLPFEDIRCGDGGKAIEATLINERGLTVDKFGLE
gi|16551957|     --------------------------------------------------
gi|7657417|      KDLTKNASVVAGTGEQCLPFDEARCGDGGKAVEATEMSEKGMASDKNGLE
gi|13649010|     KDLSKNFEVVAGTGDQCLPFEQSHCGDGGRASEASLNSERGITVDRHGFE
gi|1079143|      SQPELNWEAVVGSGERCLFGEEAHCGDGALSKDAKEAYEKGIAESSDNIE
gi|8922444|      --------------------------------------------------

                          1410      1420      1430      1440      1450
                 ....|....|....|....|....|....|....|....|....|....|
NOV4             YFVDGTMIRRIEQNGTITSLLGSNDLTS-ARELSCESVMDISQVREEWPT
gi|16551957|     --------------------------------------------------
gi|7657417|      YFVDGIMIRKVDQNGIISTLLGSNDLTS-ARELTCETSMHISQVRLEWPT
gi|13649010|     YFVDGTMIRKIDENAVITEVLGSNGLTS-TQELSCDSGMDITQVRLEWPT
gi|1079143|      YFADGTNIRMVDRDGIVSELIGNIMHKSHVKDIPCEGTIKEIEMHERWPT
gi|8922444|      --------------------------------------------------

                          1460      1470      1480      1490      1500
                 ....|....|....|....|....|....|....|....|....|....|
NOV4             DLAINPMDNSLYVLDNNYVLQISEKHQVRIVAGRPMHCQVPGIDEFLLSK
gi|16551957|     --------------------------------------------------
gi|7657417|      DLAINPMDNSIYVLDNNVVLQITEKRQVRIAAGRPMHCQVPGVEYPVG-K
```

```
gi|13649010|   DLAVNEMDNSLYVLDNNIVLQISENRRVRLIIAGRPIHCQVPGIDHFLVSK
gi|1079143|    ELAVSPMDNTLHILDDHMIERMTPLGRVRVISGRPIHCATASTAYDTD--
gi|8922444|    --------------------------------------------------

                      1510      1520      1530      1540      1550
               ....|....|....|....|....|....|....|....|....|....|
NOV4           VAIHATEESATALAVSHNCVLYIAETDEKKINRIEQVTESGEESLVACAP
gi|16551957|   --------------------------------------------------
gi|7657417|    HAVQTTLESATLLAVSYSCVLYVTETDEKKINRIRQVTIDGEISLVAAIP
gi|13649010|   VAIHSTEESARAISVSHSCLAFIAETDERKVNRIQQVTINGEIYIIACAP
gi|1079143|    LATHATEVMPQSTAFGPLCEEYVAESDSQKINRVRVIGEDERMAPFSCAE
gi|8922444|    --------------------------------------------------

                      1560      1570      1580      1590      1600
               ....|....|....|....|....|....|....|....|....|....|
NOV4           SGEDCKNDANCDCRSGDDGYEKDAKLNTPSSLAVCASGELYVADLGNIRI
gi|16551957|   --------------------------------------------------
gi|7657417|    SECDCKNDANCDCYQSGDGYEKDAKLNAPSSLAASPSGTLYIADLGNIRI
gi|13649010|   LDCQCKIDPNCDCRSGDGGYEKDAKMKAPSSLAVSPSGTLKVADLGNVRI
gi|1079143|    SKCNCLERG-CDCFRAEHYLATSANFNTIRALAVTPDSHVHIADQANYRI
gi|8922444|    --------------------------------------------------

                      1610      1620      1630      1640      1650
               ....|....|....|....|....|....|....|....|....|....|
NOV4           RFIRKNKPFLVTQNMSELSSEIDPELYLEDTTCKELYEQSLPTGDYLLNF
gi|16551957|   --------------------------------------------------
gi|7657417|    SAVSKNKPLLASMNFYEVASETDPELEIEDINCTHQYIVGVVEGDYLLNF
gi|13649010|   STVSRNQAHLNDMNIYEIASEADPELNQETVNCTELHEINLIERDYVVNF
gi|1079143|    RSVMSEIPEASPSRFYEYVAPDMPELSIENRFCQEVSERNGLLGHTTYVF
gi|8922444|    --------------------------------------------------

                      1660      1670      1680      1690      1700
               ....|....|....|....|....|....|....|....|....|....|
NOV4           RY---TGDGDSTLAITDNNSNMYNVRSDSTGMPLWEVVPDGCVYWLTHGTN
gi|16551957|   --------------------------------------------------
gi|7657417|    SN---SNDNDSTAVTDSNGNTIRRIRSDPNRMPVRYVSPDNCVIWSTIGTN
gi|13649010|   TE---NSEGDEGALESSNGNSYHERRDAGGMPLWEVVPGGCVYWLTISSN
gi|1079143|    TYNVNTSNGKDSTVIDAAGNKVFHLRDYTSQVNSIENTKGCKCRGRMTRM
gi|8922444|    --------------------------------------------------

                      1710      1720      1730      1740      1750
               ....|....|....|....|....|....|....|....|....|....|
NOV4           SAEKSVETQGHELAMMTYRCNSCLLAEESNENCWITFYEYDSFCRLENVT
gi|16551957|   --------------------------------------------------
gi|7657417|    GCEKSMTAQGLELVLFTYHCNSCLLATRSDETCWITFFDYDSECRLENVT
gi|13649010|   GVLKRNSAQGYNLALMTYPCNTCLLAEKSNENCWITVYEYDPECHLENAT
gi|1079143|    KMEHELSTPDNYNVTYEYHCPTCLLRFKLDSTCRLYVINYDEFCRLTSAV
gi|8922444|    --------------------------------------------------

                      1760      1770      1780      1790      1800
               ....|....|....|....|....|....|....|....|....|....|
NOV4           FPRCQVSSFRSDTLSSVHVQVETSSKL-DVTITTINESASGAFYTLLQLQV
```

```
gi|16551957|  --------------------------------------------------
gi|7657417|   FPTGVNTNLHGDNDKAITVDIRSSSREEDVSITTSNLSSIDSFYTMVQDQL
gi|13649010|  FPTGEVSSFHSDLEKLTKVELQTSNRE-NVLMSTNETATSTIYILKQENT
gi|1079143|   TEEGRKIELSFDLS----VKGAQVKVSENAQKEMSELIQGATVIVRNGAA
gi|8922444|   -------------MDKAITVDIRSSSREEDVSITTSNLSSIDSFYTMVQDQL

                   1810      1820      1830      1840      1850
              ....|....|....|....|....|....|....|....|....|....|
NOV4          RNSEYTIGADGSLRLLLLENGMEVALQTEEPHLEEGTVNPTVGK-------RNV
gi|16551957|  --------------------------------------------------
gi|7657417|   RNSEQIGYDGSLRITFYESGLLDSHYQTEEPHVLEGTANPTVAK------RNV
gi|13649010|  DSTYRHNPDGSLRVTPESGMEIGLSEEPHILEAGAVNPTLGK------RNV
gi|1079143|   LSRTTVDMDGSTTSITPWSHNLQMEVARYTILLAEQSELLGHSYPVPAKQR
gi|8922444|   RNSEQIGYDGSLRIIYESGLDSHYQTEEPHVLEGTANPTVAK------RNV

                   1860      1870      1880      1890      1900
              ....|....|....|....|....|....|....|....|....|....|
NOV4          ELPIQNGLNEVEWRQRKEQAR-------------SQVTVEGRRLRVENRIL
gi|16551957|  --------------------------------------------------
gi|7657417|   TEFGSNGQNLVEWRFRKEQAQ-------------GKVNVEGRKLRVKGRNL
gi|13649010|  SLPGNHNANLTEWRQRKEQNK-------------GNVSASERRLRASNENL
gi|1079143|   IEIAGDLANRFEWRYFVRKQGPLQAGKQSKGPPRPVTEVGRKERVAGDNV
gi|8922444|   ELPGENGQNLVEWRFRKEQAQ-------------GKVNVEGRKLRVKGRNL

                   1910      1920      1930      1940      1950
              ....|....|....|....|....|....|....|....|....|....|
NOV4          ESLDEDRVTPEEKIYDDHRKFTLRILEDQAERESLESESS-ELNGVNVEE
gi|16551957|  --------------------------------------------------
gi|7657417|   LSVDFDHTTKEEKIYEDHRKFLDRIAEETSGHETLKLESS-KLMAVAVEE
gi|13649010|  LSIDFDHITREGKIYDDHRKSTLRLLEDQTGREIEESEVS-EYNEVNIEE
gi|1079143|   ETLEYDREVQSVVEMVDDKQRLEAVTYDRRTSREISERDQSGDYAYVDLEE
gi|8922444|   LSVDFDRTTKEEKIYDDHRKSLERIAEETSGHETLKLESS-KLMAVNVEE

                   1960      1970      1980      1990      2000
              ....|....|....|....|....|....|....|....|....|....|
NOV4          SPGEYLTAGEQRGIMSERMEYEQAGKLTSRIEEDGKHVSYVELEKSMVLLL
gi|16551957|  --------------------------------------------------
gi|7657417|   SSTGQIASIQRGTTSEKVDYESQGRLVSRVEEDGKHVSYEYLEKSMVLLL
gi|13649010|  EPSGLVTPICRGTWNEKMEYLQSGKLIERTWEEGKNVSYEYLEKSMVLLL
gi|1079143|   DRFERLVSWKWGVLQEAYSEERNGRLNEIKYGEGSEMVYAKKQMFGSEPE
gi|8922444|   SSTGQIASEQRGTTSEKVDYEGQGRLVSRVEEDGKHVSYEYLEKSMVLLL

                   2010      2020      2030      2040      2050
              ....|....|....|....|....|....|....|....|....|....|
NOV4          HSQR----QYEFEVEKNDRESSVTRENVASQTLEEIRSVGYYRNIIQPPR
gi|16551957|  --------------------------------------------------
gi|7657417|   HSQR----QYIFEYDNWDRESAITREEVARETMQEIRSIGYYRNIYNPPR
gi|13649010|  HSQF----RYEFEYEQPECELSVEEESMVREESLGEMLSVGYYRNIYTPPD
gi|1079143|   KVTTPRRSDYILLQYDDAGAEQSLETEERGHIEAFSLQTSLGFFKYQEYSPI
gi|8922444|   HSQR----QYEFEYDMWDRESAITREESVARETMQEIRSIGYYRNIYNPPR

                   2060      2070      2080      2090      2100
```

```
              ....|....|....|....|....|....|....|....|....|....|
NOV4          GNASVIQDETEDGHLLHTFYLGTGRRVIYKYGKLSKLAETLYDTEKVSFT
gi|16551957|  --------------------------------------------------
gi|7657417|   SNASIETDYNEEGLLLQTAFLGTSRRVLFKYRRQTRLSEILYDSTRVSFT
gi|13649010|  SSTSFEQDYSRDGRLLQTLHEGTGERVLYKYTKQARLSEVLYDTEQVILT
gi|1079143|   NRHPFEILYNDECQTLAKIHPHQSGKVAFVHDTACRLETILAGLSSTHYT
gi|8922444|   SNASIETDYNEEGLLLQTAFLGTSRRVLFKYRRQTRLSEILYDSTRVSFT

                       2110      2120      2130      2140      2150
              ....|....|....|....|....|....|....|....|....|....|
NOV4          YDETAGMLKTINLQNEGFTCTIRYHQIGPLIDRQIFRFI-EEGNVNARFD
gi|16551957|  ------MLKTINLQNEGFICTIRYHQIGPLIDRQIFRFR-EEGNVNARFD
gi|7657417|   YDEIAGVLKTVNLQSDGFICTIRYHQIGPLIDRQIFRFS-EDGMVNARFD
gi|13649010|  YEESPGVIKTIHLMHEGFICTIRYHQTGPLIGRQIFRFS-EEGLVHARFD
gi|1079143|   YQDITSLVKSVEVQKPGFELRREFKYHAGLLKDEKLRFGSKNSLASARYK
gi|8922444|   YDEIAGVLKTVNLQSDGFICTIRYHQIGPLIDRQIFRFS-EDGNVNARFD

                       2160      2170      2180      2190      2200
              ....|....|....|....|....|....|....|....|....|....|
NOV4          YNYDNSFRVTSMQAVINETPLPIDLYPYDDVSGKTEQFGKFGVIYYDINQ
gi|16551957|  YNYDNSFRVTSMQAVINETPLPIDLYRYDDVSGKTEQFGKFGVIYYDINQ
gi|7657417|   YSYDNSFRVTSMQGVINETPLPIDLYQFDDISGKVEQFGKFGVIYYDINQ
gi|13649010|  YSYKN-FRVTSMQAVINHTPLPIDLYRYVDVFGKTEQFGKFSVINYDLNQ
gi|1079143|   YAYDGNAPLSGTEMAIDDKELPTTRYKESQNLGQLEVVQDLKYTRNAFNE
gi|8922444|   YSYDNSFRVTSMQGVINETPLPIDLYQFDDISGKVEQFGKFGVIYYDINQ

                       2210      2220      2230      2240      2250
              ....|....|....|....|....|....|....|....|....|....|
NOV4          IITTAVMT---HTKHFDAYGRMKEVQYEIFRSLMYWMTVQYDNMGRVVKK
gi|16551957|  IITTAVMT---HTKHFDAYGRMKEVQYEIFRSLMYWMTVQYDNMGRVVKK
gi|7657417|   IISTAVMT---YTKHFDAHGRIKEIQYEIFRSLMYWITIQYDNMGRVTKE
gi|13649010|  VITTTVMK---HTKIFSANCQVIEVQYEILKAEAYWMTIQYDNVGRMVIC
gi|1079143|   TYIQDSAKQFFAIVDYEQHGRVKSVLMNYKNIDVERLEIDYDLRNFEKSQ
gi|8922444|   IISTAVMT---YTKHFDAHGRIKEIQYEIFRSLMYWITIQYDNMGRVTKE

                       2260      2270      2280      2290      2300
              ....|....|....|....|....|....|....|....|....|....|
NOV4          ELKVGPYANTTRYSYEYLADGQLQTVSINDKPLWRYSYDLNGNLHLLSPG
gi|16551957|  ELKVGPYANTTRYSYEYLADGQLQTVSINDKPLWRYSYDLNGNLHLLSPG
gi|7657417|   EIKIGPFANTTKYAYEYLVDGQLQTVYLNEKIMWRYNYDLNGNLHLLNPS
gi|13649010|  DERVGVDANITRYFYEYLADGQLQTVSVNDRTQWRYSYDLNGNINLLSHG
gi|1079143|   KTTFGRSTAFDKINYNAEG---HVVEVLGTN-NRKYLFDENG--RTVGVV
gi|8922444|   EIKIGPFANTTKYAYEYLVDGQLQTVYLHEKIMWRYNYDLNGNLHLLNPS

                       2310      2320      2330      2340      2350
              ....|....|....|....|....|....|....|....|....|....|
NOV4          NSARLTPLRYDIRDRITRLGDVQYK-MDEDGFLRQRGGDIFEYNSAGLLI
gi|16551957|  NSARLTPLRYDIRDRITRLGDVQYK-MDEDGFLRQRGGDIFEYNSAGLLI
gi|7657417|   SSARLTPLRYDLRDRITRLGDVQYR-LDEDGFLRQRGTEIFEYSSKGLLT
gi|13649010|  KSARLTPLRYDLRDRITRLGEIQYK-MDEDGFLRQRGNDIFEYNSNGLLQ
gi|1079143|   HQGEKFNLGYLIGDRVIKVGDVEFNNYDARGFVVKRGEQKYRYNNRGQLI
gi|8922444|   NSARLTPLRYDLRDRITRLGDVQYR-LDEDGFLRQRGTEIFEYSSKGLLT
```

```
                   2360      2370      2380      2390      2400
              ....|....|....|....|....|....|....|....|....|....|
NOV4          KAYNRAGSWSVRYRYDGLGRRVSSKSSHSHHLQFFYADLTNPTKVTHLYN
gi|16551957|  KAYNRAGSWSVRYRYDGLGRRVSSKSSHSHHLQFFYADLTNPTKVTHLYN
gi|7657417|   EVYSKGSGWTVIYRYDGLGRRVSSKTSLGQHLQFFYADLTYPTRITHVYN
gi|13649010|  KAYNKASGWTVQYYYDGLGRRVASKSSLGQHLQFFYADLTNPIRVTHLYN
gi|1079143|   HSFERER-FQSWYYYDDRSRLVAWHDNKGNTTCYYYANPRTEHLVTHVHF
gi|8922444|   EVYSKGSGWTVIYRYDGLGRRVSSKTSLGQHLQFFYADLTYPTRITHVYN

                   2410      2420      2430      2440      2450
              ....|....|....|....|....|....|....|....|....|....|
NOV4          HSSSEITSLYYDLQGHLFAMEISSGDHFY-ACDNIGTPLAVFSGTGLMIK
gi|16551957|  HSSSEITSLYYDLQGHLFAMELSSGDEFYIACDNIGTPLAVFSGTGLMIK
gi|7657417|   HSSSEITSLYYDLQGHLFAMEISSGDEFYIASDNTGTPLAVFSSNGLMLK
gi|13649010|  HTSSEITSLYYDLQGHLIAMELSSGEEYYVACDNTGTPLAVFSSRGQVIK
gi|1079143|   PKISRTMKLFYDDRDMLIALEHED-QRYYVATDQNGSPLAFFDQNGSIVK
gi|8922444|   HSSSEITSLYYDLQGHLFAMEISSGDEFYIASDNTGTPLAVFSSNGLMLK

                   2460      2470      2480      2490      2500
              ....|....|....|....|....|....|....|....|....|....|
NOV4          QILYTAYGEIYMDTNPNFQLTICYHGGLYDPLTKLVIMGRRDYDVLAGRW
gi|16551957|  QILYTAYGEIYMDTNPNFQLTICYHGGLYDPLTKLVHMGRRDYDVLAGRW
gi|7657417|   QTQYTAYGEIYFDSNVDFQLVIGFHGGLYDPLTKLIHFGERDYDILAGRW
gi|13649010|  EILYTPYGDIYHDTYPDFQVIIGFHGGLYDFLTKLVHLGQRDYDVYAGRW
gi|1079143|   EMKRIPFGRIIKDEKFEFFVPEDFHGGLIDPHTKLVVTEQRQYDPHVGQW
gi|8922444|   QLQYTAYGEIYFDSNIDFQLVIGFHGGLYDPLTKLIHFGERDYDILAGRW

                   2510      2520      2530      2540      2550
              ....|....|....|....|....|....|....|....|....|....|
NOV4          TSPDHELWKHISSSNVMPFNLYMFKNNIPISNSQDIKCFMTDVNSWLLTF
gi|16551957|  TSPDHELWKHGSSSNVMPFNLYMFKANHIPISNSQDIKCFMTDVNSWLLTF
gi|7657417|   TTPDIEIWKRLGK-DFAPFNLYMFRNNNPASKIHDVKDYITDVNSWLVTF
gi|13649010|  TTPNHHIWKQTNL-LEKFFHLYSFENNYPVGKIQDVAKYTTDHRSWLRLF
gi|1079143|   MEPLWETLATEMS-HETDVFITRYHNNQPINP-NKPQNYMIDLDSWLQLF
gi|8922444|   TTPDIEIWKRIGK-DFAPFNLYMFRNNNPASKIHDVKDYITDVNSWLVTF

                   2560      2570      2580      2590      2600
              ....|....|....|....|....|....|....|....|....|....|
NOV4          GFQLHNVEPGYEKPDYDAMEPSYELIHTQMKTQEWDNSKSHLGVQCEVQK
gi|16551957|  GFQLHNVEPGYEKPDMDAMEPSYELIHTQMKTQEWDNSKSHLGVQCEVQK
gi|7657417|   GFHLHNAIPGFEVPKFDLTEPSYELVK----SQQWEDVPPIFGVCQQVAK
gi|13649010|  GFQLHNVLPGFLKPELENLFLTYELLRLQTKTQEWDPGKTHLGICCELQK
gi|1079143|   SYDLNNMQS-SRYTKLAQYTEQASTKS----NTLAPDFGVISGLECIVSK
gi|8922444|   SFHLHNAIPGFEVPKFDLTEPSYELVK----SQQWDDIPPIFGVCQQVAK

                   2610      2620      2630      2640      2650
              ....|....|....|....|....|....|....|....|....|....|
NOV4          QLKAFVILERFDQLYGSTITSCQQAPKTKKFASSGSVFGKGVKFALKDGR
gi|16551957|  QLKAFVILERFDQLYGSTITSCLQAPKTKKFASSGSVFGKGVKFALKDGR
gi|7657417|   QAKAFLSIGKMAEVQVSRRK-AGAEQSWLWEATVKSLISKGVMLAVSQGR
gi|13649010|  QLSNFISLDQLPMTPRYNDGRCLEGGKQPREAAVPSVFGKGIKFALKDGI
```

```
gi|1079143|     TSEKESDFDFVPKPLLKTEP--KMRNLLPRVSYRRGVEGEGVLLSRIGGR
gi|8922444|     QAKAFLSLGKNAKVQVSRRR-AGGAQSWLWFATVKSLTGKGVMLAVSQGR

                        2660      2670      2680      2690      2700
                ....|....|....|....|....|....|....|....|....|....|
NOV4            VTSDIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFVKPGPSEGD
gi|16551957|    VTTDIISVANEDGRRVAAILNHAHYLENLHFTIDGVDTHYFVKPGPSEGD
gi|7657417|     VQTNVLNIANEDCIKVAAVLNNAFYLENLHFTIEGKDTHYFIKTTTPESD
gi|13649010|    VTADTIGVANEDSRRIAATLNNAHYLENLHFTIEGRDTHYFIKLGSLEED
gi|1079143|     ALVSVVDGSNSVVQDVVSSVFNNSYFLDLHFSIHDQDVFYFVKDN-----
gi|8922444|     VQTNVLNIANEDCIKVAAVLNNAFYLENLHFTIEGKDTHYFIKTTTPESD

                        2710      2720      2730      2740      2750
                ....|....|....|....|....|....|....|....|....|....|
NOV4            LAILGLSGGRKTLENGVNVTVSQINTVLNGRTRRYTDIQLQYGALCLNTR
gi|16551957|    LAILGLSGGRKTLENGVNVTVSQINTVLSGRTRRYTDIQLQYGALCLNTR
gi|7657417|     LGTLRLTSGRKALENGINVTVSQSTTVVNGRTPRFADVEMQFGALALHVR
gi|13649010|    LVLIGNLGGRRILENGVNVTVSQMTSVLNGRTPRFADIQLQHGALCFNIR
gi|1079143|     --VLKLRDDNEELRRLGGNFNISTHEISDHGGSAAKELRLHGPDAVLIEK
gi|8922444|     LGTLRLTSGRKALENGINVTVSQSTTVVNGRTPRFADVSMQFGALALHVR

                        2760      2770      2780      2790      2800
                ....|....|....|....|....|....|....|....|....|....|
NOV4            YGTTLDEEKARVLELARQRAVRQAVAREQQRLPEGEEGLPAWTEGEKQQV
gi|16551957|    YGTTLDEEKARVLELARQRAVRQAVAREQQRLPEGEEGLPAWTEGEKQQV
gi|7657417|     YGMTLDEEKARILEQARQRALARAVAREQQFVRDGEEGARLWTEGEKRQL
gi|13649010|    YGTTVPEEKNHVLEIARQRAVAQAWTKEQRRLQPGEEGIRAWTEGEKQQL
gi|1079143|     YGVDPEQERHRILKHPHKRAVERAWELEKQIVAAGFQGRGDWTEEERHEL
gi|8922444|     YGMTLDEEKARILEQARQRALARAVAREQQRVRDGEEGARLWTEGEKRQL

                        2810      2820      2830      2840      2850
                ....|....|....|....|....|....|....|....|....|....|
NOV4            LSTGRVQGYDGFEVILSVEQYPE.SLSANNIHFNRQSHMGRR---------
gi|16551957|    LSTGRVQGYDGFEVILSVEQYPELSDSANNIHFMRQSEMGRR---------
gi|7657417|     LSAGKVQGYDGYYVVLSVEQYPELADSANNIQFLRQSEIGKR---------
gi|13649010|    LSTGRVQGYDGYFVVLSVEQYLELSDSANNIHFMRQSEIGRR---------
gi|1079143|     VQHGDVDEWNGIDEHSRHKYPQLADDPGNVAEQPDAKRKRRKTGSSHRSA
gi|8922444|     LSAGKVQGYDGYYVVLSVEQYPELADSANNIQFLRQSEIGRR---------

                        2860
                ....|....|...
NOV4            -------------
gi|16551957|    -------------
gi|7657417|     -------------
gi|13649010|    -------------
gi|1079143|     SNRRQLKFGELSA
gi|8922444|     -------------
```

Table 4E. Domain Analysis of NOV4

gnl|Pfam|pfam01500, Keratin_B2, Keratin, high sulfur B2 protein. High sulfur proteins are cysteine-rich proteins synthesized during the differentiation of hair matrix cells, and form hair fibers in association with hair keratin intermediate filaments. This family has been divided up into four regions, with the second region containing 8 copies of a short repeat. This family is also known as B2 or KAP1.

CD-Length = 144 residues, 87.5% aligned
Score = 38.9 bits (89), Expect = 0.004

```
Query:  630   CIDVACSNHGTCITGTCICNPGYKGESCEEVDCMDPTCSGRGVCVRGECHCFVGWGGTNC
689
              C    CS  GTC +  C      +  SC +  C  P CS    C R  C        + C
Sbjct:  5     CGFPTCSTLGTCGSSCC------QPPSCCQPSCCQPVCSQTTCC-RPTCFQSSCCRPSCC
57

Query:  690   ETP--RATCLDQCSGHGTFLPDTGLCSCDPSWTGHDCSIEICAADCGGHGVCVGGTCRCE
747
              +T   + TC    S         G+ SC   W   DC +E
Sbjct:  58    QTSCCQPTCCQSSSCQ----TGCGIGSCRTRWCRPDCRVE--------------------
93

Query:  748   DGWMGAACDQRACHPRCAEHGTCRDGKCECS---PGWNGEHC  786
                     C    C   C     C+    + S   P + G+ C
Sbjct:  94    -----GTCLPPCCVVSCTPPTCCQPVSAQASCCRPSYCGQSC  130
```

The novel TEN-M-like protein encoded by the gene of invention has highest homology to the mouse TEN-M4 protein, which belongs to the ODZ/TENM family of proteins. This family was first identified in Drosophila as being a pair-rule gene affecting segmentation of the early embryo. It was the first pair-rule gene identified that was not a transcription factor, but a type II transmembrane protein. Vertebrate homologs of the TENM family have been identified in mouse and zebrafish. In the mouse, TEN-M4 expression was found to be on the cell surface, in the brain, trachea as well as developing limb and bone. Analysis of the TEN-M1 protein reveals that it can bind to itself, making it likely that TEN-M4 may be a dimeric moiety as well. In cell culture experiments, fragments of the TEN-M proteins can bind the Drosophila PS2 integrins. In addition, members of the TEN-M family have been identified to be downstream of the endoplasmic reticulum stress response pathway, which alters the response of cells to their environment. This suggests that the ODZ/TENM family may be involved in cell adhesion, spreading and motility. Translocations leading to the fusion of this gene with the NRG1/HGL gene from chromosome 8 have been found to generate a paracrine growth factor for one mammary carcinoma cell line, termed gamma-heregulin. Therefore this novel gene may have widespread implications in development, regeneration and carcinogenesis of various tissues.

Two new potential ligands of the Drosophila PS2 integrins have been characterized by functional interaction in cell culture. These potential ligands are a new Drosophila laminin alpha2 chain encoded by the wing blister locus and Ten-m, an extracellular protein known to be involved in embryonic pattern formation. As with previously identified PS2 ligands, both contain RGD sequences, and RGD-containing fragments of these two proteins (DLAM-RGD and TENM-RGD) can support PS2 integrin-mediated cell spreading. In all cases, this spreading is inhibited specifically by short RGD-containing peptides. As previously found for the PS2 ligand tiggrin (and the tiggrin fragment TIG-RGD), TENM-RGD induces maximal spreading of cells expressing integrin containing the alphaPS2C splice variant. This is in contrast to DLAM-RGD, which is the first Drosophila polypeptide shown to interact preferentially with cells expressing the alphaPS2 m8 splice variant. The betaPS integrin subunit also varies in the presumed ligand binding region as a result of alternative splicing. For TIG-RGD and TENM-RGD, the beta splice variant has little effect, but for DLAM-RGD, maximal cell spreading is supported only by the betaPS4A form of the protein. Thus, the diversity in PS2 integrins due to splicing variations, in combination with diversity of matrix ligands, can greatly enhance the functional complexity of PS2-ligand interactions in the developing animal. The data also suggest that the splice variants may alter regions of the subunits that are directly involved in ligand interactions, and this is discussed with respect to models of integrin structure.

A sequence of about thirty to forty amino-acid residues long found in the sequence of epidermal growth factor (EGF) has been shown to be present, in a more or less conserved form, in a large number of other, mostly animal proteins. The list of proteins currently known to contain one or more copies of an EGF-like pattern is large and varied. The functional significance of EGF domains in what appear to be unrelated proteins is not yet clear. However, a common feature is that these repeats are found in the extracellular domain of membrane-bound proteins or in proteins known to be secreted (exception: prostaglandin G/H synthase). The EGF domain includes six cysteine residues which have been shown (in EGF) to be involved in disulfide bonds. The main structure is a two-stranded beta-sheet followed by a loop to a C-terminal short two-stranded sheet. Subdomains between the conserved cysteines vary in length. The NHL (NCL-1, HT2A and LIN-41) repeat is found in a variety of enzymes of the copper type II, ascorbate-dependent monooxygenase family which catalyse the C-terminus alpha-amidation of biological peptides. The repeat also occurs in a human zinc finger protein that specifically interacts with the activation domain of lentiviral Tat proteins. The repeat domain that is often associated with RING finger and B-box motifs (see, Ben-Zur T,

Dev Biol 2000 Jan 1;217(1):107-20; Adelaide J, Int J Oncol 2000 Apr;16(4):683-8 ; Wang XZ, Oncogene 1999 Oct 7;18(41):5718-21; Schaefer G, Oncogene 1997 Sep 18;15(12):1385-94 ; Wang XZ, EMBO J 1998 Jul 1;17(13):3619-30; Baumgartner S, EMBO J 1994 Aug 15;13(16):3728-40; Otaki JM, Dev Biol 1999 Aug 1;212(1):165-81; Mieda M, Mech Dev 1999 Sep;87(1-2):223-7; Oohashi T, J Cell Biol 1999 May 3;145(3):563-77; Graner MW, J Biol Chem 1998 Jul 17;273(29):18235-41, incorporated herein by reference).

The protein similarity information, expression pattern, and map location for the TEN-M4-like protein and nucleic acid disclosed herein suggest that this TEN-M4-like protein may have important structural and/or physiological functions characteristic of this family. Therefore, the nucleic acids and proteins of the invention are useful in potential diagnostic and therapeutic applications and as a research tool. These include serving as a specific or selective nucleic acid or protein diagnostic and/or prognostic marker, wherein the presence or amount of the nucleic acid or the protein are to be assessed, as well as potential therapeutic applications such as the following: (i) a protein therapeutic, (ii) a small molecule drug target, (iii) an antibody target (therapeutic, diagnostic, drug targeting/cytotoxic antibody), (iv) a nucleic acid useful in gene therapy (gene delivery/gene ablation), and (v) a composition promoting tissue regeneration *in vitro* and *in vivo* (vi) biological defense weapon.

The NOV4 nucleic acids and proteins of the invention are useful in potential diagnostic and therapeutic applications implicated in various diseases and disorders described below and/or other pathologies. For example, the compositions of the present invention will have efficacy for treatment of patients suffering from: cardiac diseases, myocardial contractility in failing heart and other diseases, disorders and conditions of the like. The disclosed NOV4 nucleic acid of the invention encoding a TEN-M4-like protein includes the nucleic acid whose sequence is provided in Table 4A or a fragment thereof. The invention also includes a mutant or variant nucleic acid any of whose bases may be changed from the corresponding base shown in Table 4A while still encoding a protein that maintains TEN-M4-like protein-like activities and physiological functions, or a fragment of such a nucleic acid. The invention further includes nucleic acids whose sequences are complementary to those just described, including nucleic acid fragments that are complementary to any of the nucleic acids just described. The invention additionally includes nucleic acids or nucleic acid fragments, or complements thereto, whose structures include chemical modifications. Such modifications include, by way of nonlimiting example, modified bases, and nucleic acids whose sugar phosphate backbones are modified or derivatized. These modifications are carried out at least in part to enhance the chemical stability of the modified nucleic acid, such that they may be

used, for example, as antisense binding nucleic acids in therapeutic applications in a subject. In the mutant or variant nucleic acids, and their complements, up to about 11 percent of the bases may be so changed.

The disclosed NOV4 protein of the invention includes the TEN-M4-like protein whose sequence is provided in Table 3B. The invention also includes a mutant or variant protein any of whose residues may be changed from the corresponding residue shown in Table 4B while still encoding a protein that maintains beta adrenergic receptor kinase-like activities and physiological functions, or a functional fragment thereof. In the mutant or variant protein, up to about 3 percent of the residues may be so changed.

The protein similarity information, expression pattern, and map location for TEN-M4-like protein and nucleic acid (NOV4) disclosed herein suggest that NOV4 may have important structural and/or physiological functions characteristic of the TEN-M4 protein family. Therefore, the NOV4 nucleic acids and proteins of the invention are useful in potential diagnostic and therapeutic applications. These include serving as a specific or selective nucleic acid or protein diagnostic and/or prognostic marker, wherein the presence or amount of the nucleic acid or the protein are to be assessed, as well as potential therapeutic applications such as the following: (i) a protein therapeutic, (ii) a small molecule drug target, (iii) an antibody target (therapeutic, diagnostic, drug targeting/cytotoxic antibody), (iv) a nucleic acid useful in gene therapy (gene delivery/gene ablation), and (v) a composition promoting tissue regeneration *in vitro* and *in vivo*.

The NOV4 nucleic acids and proteins of the invention are useful in potential diagnostic and therapeutic applications implicated in various diseases and disorders described below. For example, the compositions of the present invention will have efficacy for treatment of patients suffering from: Von Hippel-Lindau (VHL) syndrome, Alzheimer's disease, stroke, tuberous sclerosis, hypocalcaemia, Parkinson's disease, Huntington's disease, cerebral palsy, epilepsy, Lesch-Nyhan syndrome, multiple sclerosis, ataxia-telangiectasia, leukodystrophies, behavioral disorders, addiction, anxiety, pain, neurodegeneration, fertility disorders, hyperparathyroidism, hypoparathyroidism, cardiomyopathy, atherosclerosis, hypertension, congenital heart defects, aortic stenosis, atrial septal defect (ASD), atrioventricular (A-V) canal defect, ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular septal defect (VSD), valve diseases, tuberous sclerosis, scleroderma, obesity, transplantation disorders, diabetes, autoimmune disease, renal artery stenosis, interstitial nephritis, glomerulonephritis, polycystic kidney disease, systemic lupus erythematosus, renal tubular acidosis, IgA nephropathy, hypocalcaemia, asthma, emphysema, scleroderma, allergy, ARDS, Hirschsprung's disease,

WHAT IS CLAIMED IS:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:
    - (a) a mature form of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, and 34;
    - (b) a variant of a mature form of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, and 34, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of the amino acid residues from the amino acid sequence of said mature form;
    - (c) an amino acid sequence selected from the group consisting SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, and 34; and
    - (d) a variant of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, and 34, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of amino acid residues from said amino acid sequence.

2 The polypeptide of claim 1, wherein said polypeptide comprises the amino acid sequence of a naturally-occurring allelic variant of an amino acid sequence selected from the group consisting SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, and 34.

3. The polypeptide of claim 2, wherein said allelic variant comprises an amino acid sequence that is the translation of a nucleic acid sequence differing by a single nucleotide from a nucleic acid sequence selected from the group consisting of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, and 35.

4. The polypeptide of claim 1, wherein the amino acid sequence of said variant comprises a conservative amino acid substitution.

5. An isolated nucleic acid molecule comprising a nucleic acid sequence encoding a polypeptide comprising an amino acid sequence selected from the group consisting of:
   - (a) a mature form of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, and 34;
   - (b) a variant of a mature form of an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, and 34, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of the amino acid residues from the amino acid sequence of said mature form;
   - (c) an amino acid sequence selected from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, and 34;
   - (d) a variant of an amino acid sequence selected from the group consisting SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, and 34, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of amino acid residues from said amino acid sequence;
   - (e) a nucleic acid fragment encoding at least a portion of a polypeptide comprising an amino acid sequence chosen from the group consisting of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, and 34, or a variant of said polypeptide, wherein one or more amino acid residues in said variant differs from the amino acid sequence of said mature form, provided that said variant differs in no more than 15% of amino acid residues from said amino acid sequence; and
   - (f) a nucleic acid molecule comprising the complement of (a), (b), (c), (d) or (e).

6. The nucleic acid molecule of claim 5, wherein the nucleic acid molecule comprises the nucleotide sequence of a naturally-occurring allelic nucleic acid variant.

7. The nucleic acid molecule of claim 5, wherein the nucleic acid molecule encodes a polypeptide comprising the amino acid sequence of a naturally-occurring polypeptide variant.

8. The nucleic acid molecule of claim 5, wherein the nucleic acid molecule differs by a single nucleotide from a nucleic acid sequence selected from the group consisting of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, and 35.

9. The nucleic acid molecule of claim 5, wherein said nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of:
   (a) a nucleotide sequence selected from the group consisting of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, and 35;
   (b) a nucleotide sequence differing by one or more nucleotides from a nucleotide sequence selected from the group consisting of SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, and 35, provided that no more than 20% of the nucleotides differ from said nucleotide sequence;
   (c) a nucleic acid fragment of (a); and
   (d) a nucleic acid fragment of (b).

10. The nucleic acid molecule of claim 5, wherein said nucleic acid molecule hybridizes under stringent conditions to a nucleotide sequence chosen from the group consisting SEQ ID NOS:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, and 35, or a complement of said nucleotide sequence.

11. The nucleic acid molecule of claim 5, wherein the nucleic acid molecule comprises a nucleotide sequence selected from the group consisting of:
   (a) a first nucleotide sequence comprising a coding sequence differing by one or more nucleotide sequences from a coding sequence encoding said amino acid sequence, provided that no more than 20% of the nucleotides in the coding sequence in said first nucleotide sequence differ from said coding sequence;
   (b) an isolated second polynucleotide that is a complement of the first polynucleotide; and
   (c) a nucleic acid fragment of (a) or (b).

12. A vector comprising the nucleic acid molecule of claim 11.

13. The vector of claim 12, further comprising a promoter operably-linked to said nucleic acid molecule.

14. A cell comprising the vector of claim 12.

15. An antibody that binds immunospecifically to the polypeptide of claim 1.

16. The antibody of claim 15, wherein said antibody is a monoclonal antibody.

17. The antibody of claim 15, wherein the antibody is a humanized antibody.

18. A method for determining the presence or amount of the polypeptide of claim 1 in a sample, the method comprising:
    (a) providing the sample;
    (b) contacting the sample with an antibody that binds immunospecifically to the polypeptide; and
    (c) determining the presence or amount of antibody bound to said polypeptide,

thereby determining the presence or amount of polypeptide in said sample.

19. A method for determining the presence or amount of the nucleic acid molecule of claim 5 in a sample, the method comprising:
    (a) providing the sample;
    (b) contacting the sample with a probe that binds to said nucleic acid molecule; and
    (c) determining the presence or amount of the probe bound to said nucleic acid molecule,

thereby determining the presence or amount of the nucleic acid molecule in said sample.

20. The method of claim 19 wherein presence or amount of the nucleic acid molecule is used as a marker for cell or tissue type.

21. The method of claim 20 wherein the cell or tissue type is cancerous.

22. A method of identifying an agent that binds to a polypeptide of claim 1, the method comprising:
    (a) contacting said polypeptide with said agent; and
    (b) determining whether said agent binds to said polypeptide.

23. The method of claim 22 wherein the agent is a cellular receptor or a downstream effector.

24. A method for identifying an agent that modulates the expression or activity of the polypeptide of claim 1, the method comprising:
    (a) providing a cell expressing said polypeptide;
    (b) contacting the cell with said agent, and
    (c) determining whether the agent modulates expression or activity of said polypeptide,

whereby an alteration in expression or activity of said peptide indicates said agent modulates expression or activity of said polypeptide.

25. A method for modulating the activity of the polypeptide of claim 1, the method comprising contacting a cell sample expressing the polypeptide of said claim with a compound that binds to said polypeptide in an amount sufficient to modulate the activity of the polypeptide.

26. A method of treating or preventing a NOVX-associated disorder, said method comprising administering to a subject in which such treatment or prevention is desired the polypeptide of claim 1 in an amount sufficient to treat or prevent said NOVX-associated disorder in said subject.

27. The method of claim 26 wherein the disorder is selected from the group consisting of cardiomyopathy and atherosclerosis.

28. The method of claim 26 wherein the disorder is related to cell signal processing and metabolic pathway modulation.

29. The method of claim 26, wherein said subject is a human.

30. A method of treating or preventing a NOVX-associated disorder, said method comprising administering to a subject in which such treatment or prevention is desired

the nucleic acid of claim 5 in an amount sufficient to treat or prevent said NOVX-associated disorder in said subject.

31. The method of claim 30 wherein the disorder is selected from the group consisting of cardiomyopathy and atherosclerosis.

32. The method of claim 30 wherein the disorder is related to cell signal processing and metabolic pathway modulation.

33. The method of claim 30, wherein said subject is a human.

34. A method of treating or preventing a NOVX-associated disorder, said method comprising administering to a subject in which such treatment or prevention is desired the antibody of claim 15 in an amount sufficient to treat or prevent said NOVX-associated disorder in said subject.

35. The method of claim 34 wherein the disorder is diabetes.

36. The method of claim 34 wherein the disorder is related to cell signal processing and metabolic pathway modulation.

37. The method of claim 34, wherein the subject is a human.

38. A pharmaceutical composition comprising the polypeptide of claim 1 and a pharmaceutically-acceptable carrier.

39. A pharmaceutical composition comprising the nucleic acid molecule of claim 5 and a pharmaceutically-acceptable carrier.

40. A pharmaceutical composition comprising the antibody of claim 15 and a pharmaceutically-acceptable carrier.

41. A kit comprising in one or more containers, the pharmaceutical composition of claim 38.

42. A kit comprising in one or more containers, the pharmaceutical composition of claim 39.

43. A kit comprising in one or more containers, the pharmaceutical composition of claim 40.

44. A method for determining the presence of or predisposition to a disease associated with altered levels of the polypeptide of claim 1 in a first mammalian subject, the method comprising:
   (a) measuring the level of expression of the polypeptide in a sample from the first mammalian subject; and
   (b) comparing the amount of said polypeptide in the sample of step (a) to the amount of the polypeptide present in a control sample from a second mammalian subject known not to have, or not to be predisposed to, said disease;

wherein an alteration in the expression level of the polypeptide in the first subject as compared to the control sample indicates the presence of or predisposition to said disease.

45. The method of claim 44 wherein the predisposition is to a cancer.

46. A method for determining the presence of or predisposition to a disease associated with altered levels of the nucleic acid molecule of claim 5 in a first mammalian subject, the method comprising:
   (a) measuring the amount of the nucleic acid in a sample from the first mammalian subject; and
   (b) comparing the amount of said nucleic acid in the sample of step (a) to the amount of the nucleic acid present in a control sample from a second mammalian subject known not to have or not be predisposed to, the disease;

wherein an alteration in the level of the nucleic acid in the first subject as compared to the control sample indicates the presence of or predisposition to the disease.

47. The method of claim 46 wherein the predisposition is to a cancer.

48. A method of treating a pathological state in a mammal, the method comprising administering to the mammal a polypeptide in an amount that is sufficient to alleviate the pathological state, wherein the polypeptide is a polypeptide having an amino acid sequence at least 95% identical to a polypeptide comprising an amino acid sequence of at least one of SEQ ID NOS:2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, and 34, or a biologically active fragment thereof.

49. A method of treating a pathological state in a mammal, the method comprising administering to the mammal the antibody of claim 15 in an amount sufficient to alleviate the pathological state.